Weather Radar Information and Distributed Hydrological Modelling (Proceedings of symposium HS03 held during IUGG2003 at Sapporo, July 2003). IAHS Publ. no. 282, 2003.

Combining physically-based and conceptual approaches in the development and parameterization of a distributed system

VICTOR KOREN, SEANN REED, MICHAEL SMITH & ZIYA ZHANG

Office of Hydrologic Development, National Weather Service, 1325 East-West Highway, Silver Spring, Maryland 20910, USA

victor.koren@noaa.gov

Abstract A grid-based Hydrology Laboratory Research Modelling System (HL-RMS) that combines lumped conceptual and distributed model features has been developed by the US National Weather Service (NWS). HL-RMS consists of a well-tested conceptual water balance model applied on a regular spatial grid and physically-based kinematic routing models. A parameter estimation procedure that combines spatially-distributed and "integrated" at basin outlets properties is discussed. Initial tests using high-resolution radar precipitation estimates show that HL-RMS yields results comparable to well-calibrated lumped model simulations and outperforms a lumped model over basins where rainfall variability effects are significant. Combining outlet information and spatially-variable basin properties provides a reliable procedure to estimate distributed model parameters. While uniform adjustment of water balance model parameters provides reasonable results at basin outlets, the problem of removing scale effects in nested basins is still a challenge.

Key words distributed model; kinematic routing; radar rainfall estimates; scaling effects; spatially-variable parameters; water balance

INTRODUCTION

The growing availability of radar precipitation estimates and other sets of spatial data have intensified research in distributed modelling. Numerous distributed models and modelling approaches have been developed to address a wide variety of issues. Models ranging in complexity from the so-called "physically-based fully distributed" to "semi-distributed" or conceptual lumped models applied at smaller scales have been constructed. The range of issues that must be addressed in distributed modelling can be found in the most recent special issue of *Hydrological Processes* (Beven & Feyen, 2002), a collection of more than 20 papers. As Beven stated in the preface to the issue, "if the reader perceives that this issue reveal the problems to be addressed in the application of distributed models more than providing solutions, then that properly reflects the current state of the art". The main problem is how to apply point process laws to the basin scale given the tremendous vertical and horizontal heterogeneities of basin properties. Beven (1995) argues that the aggregation approach towards large-scale hydrological modelling is an inadequate approach to the scale problem. As a result, parameter estimation procedures for physically-based models are not well defined, and the physically-based model may be essentially reverting to a type of elegant black box (Loague, 1990). Vieux & Moreda (2002) suggested a manual "ordered physics-based parameter adjustment" procedure assuming that there are reasonable *a priori* estimates of a limited number of spatially-variable model parameters.

In light of these modelling trends, concerns, and the specific needs of the US National Weather Service (NWS), the use of simpler approaches was investigated. This study echoes the sentiment of Robinson & Sivapalan (1995) who stated that there are not enough analyses in finding connections between physically-based and conceptual models, "although this is precisely what is required for the advancement of hydrological modelling for predictive purposes". We believe that more comprehensive analysis of conceptual lumped and physically-based distributed models on large basins is needed to fully benefit from existing lumped modelling experience. In this paper we present a preliminary attempt to combine lumped conceptual and distributed model features in the development and parameterization of the Hydrology Laboratory Research Modelling System (HL-RMS).

HL-RMS STRUCTURE

HL-RMS is a flexible modelling system, able to use grid cells or sub-basins as the computational elements for rainfall–runoff modelling. Currently, HL-RMS is defined on a regular rectangular grid. Each grid cell consists of a water balance component and a hillslope and channel routing component. A number of conceptual hillslopes at each grid cell are defined to make overland flow distances physically realistic for the relatively large cell size (16 km^2). A drainage density parameter is used to subdivide a cell into equally sized overland flow planes. Conceptual hillslopes drain water to a conceptual channel within the same grid cell. A conceptual channel usually represents the highest order stream of a selected grid cell. It is assumed that all hillslopes have the same properties inside each grid cell but they may be different from cell to cell. Cell-to-cell channel routing is done using a flow direction grid. An automatic procedure was developed to generate a coarser resolution flow direction grid from higher resolution DEM data. To facilitate efficient routing calculations, the drainage network is translated into a computational sequence of grid cells in an upstream to downstream order.

Fast response runoff from the water balance model is routed over conceptual hillslopes within each cell to a conceptual channel. Slow response runoff is assumed to enter the channel system directly from the soil and therefore bypass the hillslope routing. There is no physical connection between soil moisture states in adjacent grid cells. The conceptual channel is the only source of water exchange between neighboring pixels.

The water balance component of the current version of HL-RMS uses the Sacramento Soil Moisture Accounting Model (SAC-SMA) (Burnash, 1995), and hillslope-channel routing employs the kinematic wave model. Several factors played a role in this selection. Use of the SAC-SMA is a practical choice because NWS hydrologists have great experience with lumped applications of the model. Also, the work of Koren *et al.* (2000) established relationships between SAC-SMA parameters and soil properties, making it possible to run simulations using parameter estimates that vary within a basin. The kinematic wave model is well tested, and it provides reasonable accuracy under a wide range of flow conditions.

The SAC-SMA represents conceptually heterogeneous runoff processes over a range of spatial scales from tens to a few thousand square kilometres. There are strong physical arguments and application results to support the model structure. The basic design of SAC-SMA centres on a two layer structure: a relatively thin upper layer, and usually a much thicker lower layer which supplies moisture to meet the evapo-transpiration demands. Each layer consists of tension and free water storages that interact to generate soil moisture states and six runoff components. Partitioning of rainfall into surface runoff and infiltration is governed by the upper layer soil moisture deficit and a percolation rate to the lower layer.

A fairly general numerical scheme (Smith, 1980) is used to integrate kinematic wave equations. It employs weighting factors α and β in the spatial and temporal dimensions, respectively. The scheme stability ranges from unconditionally stable ($\alpha = 0$ and $\beta = 1$) to unconditionally unstable ($\alpha > \beta$). The best approximation is obtained at $\alpha = \beta = 0.5$, when the scheme has second order accuracy. However, it is practically impossible to satisfy the scheme stability criteria (the Courant number $C_r = 1$) at this value of weighting factors under variable channel properties. Because HL-RMS is applied over a wide range of hillslope-channel hydraulic conditions, values of 0 and 1 are used for temporal and spatial weight factors, respectively, to provide the unconditional stability. While this will lead to some reduction of simulation accuracy, our tests suggest that accuracy degradation is less significant than input data and model/parameter uncertainties for basins studied. It is also important that truncation errors of the scheme increase independently of the space–time increments ratio as α and β depart from the value of 0.5 that allows a flexible selection of space–time increments to compensate some accuracy reduction.

HL-RMS PARAMETERIZATION

Water balance and routing model parameters are assumed to be constant within each grid cell; however, they can vary from cell to cell. The approach adopted here is a two-step procedure: derivation of *a priori* parameter grids, and adjustment of these grids using observed outlet hydrographs. The basic idea used to derive parameter grids is to combine distributed grid cell data (e.g. slope, soil properties) with integrated basin properties observed at the outlet (e.g. flow measurement information, discharge time series). While this approach was applied in the first step of the routing parameter estimation procedure, it was used only in the second step of water balance model parameterization.

To account for the spatial variability within a basin, *a priori* SAC-SMA parameter grids developed by Koren *et al.* (2000) were used. High-resolution soil data (texture, hydrological group, and depth) for the conterminous US were applied. Results from lumped simulations using basin averaged *a priori* parameters (Koren *et al.*, 2000, 2002) suggest that while *a priori* estimates cannot outperform results from well calibrated parameters on gauged basins, these values provide reasonable initial estimates for ungauged basins. To improve the initial estimates, HL-RMS scales *a priori* grids over selected basin by the ratio of the SAC-SMA parameter from lumped calibration to the average parameter value from the *a priori* grid (Koren *et al.*, 2002). Because lumped-calibrated SAC-SMA parameters are scale-dependent (Koren *et al.*, 1999), some further fine tuning of selected scaled parameters might be required in distributed simulations. Fortunately, the scale effect is reduced significantly if a ratio of calibration and distributed simulation scales decreases. Koren *et al.* (1999) observed about a 3% runoff difference when calibrated SAC-SMA parameters over 1000 km^2 basin were applied at a grid scale of 16 km^2.

The kinematic wave relationship between discharge and channel cross-section was approximated by a power function defined by two parameters, specific discharge, Q_s, and exponent value, m. Two options have been developed to define these parameters: (a) use of the Chezy–Manning approximation assuming a prismatic channel, referred as a channel parameterization method, and (b) direct estimation of these parameters from an empirical discharge and cross-section relationship, referred to as a rating curve method. The basic idea for both options is to disaggregate information from outlet measurements representing basin-integrated properties into interior grid cell parameters using local geomorphologic properties. Channel shape parameters (channel parameterization option) or specific discharge and exponent (rating curve option) are first estimated at the basin outlet by fitting a curve to a plot of cross-section *vs* top width data, or to a plot of cross-section *vs* discharge data, respectively.

Two geomorphic assumptions that follow from channel geometry laws are used to estimate channel shape parameters at upstream cells: (a) the ratio of channel-forming flows at different cells, $r_{Q,i}$, equals the ratio of drainage areas, F, and (b) the ratio of cross-sectional areas of different channels, $r_{A,i}$, is a known function of stream orders. These assumptions lead to simple estimation procedures. In the channel parameterization method, when a two-parametric power approximation of channel cross-section was used, distributed channel top width parameter, a_i, can be estimated from:

$$a_i = (b+1)^{-b}\left[\frac{\sqrt{S_i}}{n_i}\frac{(r_{A,i}A_o)^m}{r_{Q,i}Q_o}\right]^{3(b+1)/2} \tag{1}$$

where S_i and n_i are channel slope and roughness, Q_o and A_o are discharge and cross-section at the outlet, and b is a shape parameter assumed to be constant within a basin, equaling an estimated value at the outlet. With known channel geometry at each cell, values of channel specific discharge, Q_s, and exponent parameter m are estimated from:

$$Q_{s,i} = \frac{\sqrt{S_i}}{n_i}\left[a_i(b+1)^b\right]^{-2/3(b+1)}, \quad \text{and} \quad m = \frac{b+5/3}{b+1} \tag{2}$$

In the rating curve method, specific discharge at each grid cell is estimated from:

$$Q_{s,i} = Q_{s,o}\frac{F_i}{F_o}{r_{A,i}}^m \tag{3}$$

As in the channel parameterization method, the exponent parameter m is a constant that is equal to an estimated outlet value. An advantage of this method is that it does not require estimation of channel slope and roughness at upstream points. However, because of this, the method may poorly represent channel geometry if there are distinct portions of basin, which have significant differences in average channel slopes (e.g. transition from mountain region to a plain). Note that values of the specific discharge and exponent parameters from these two options may differ because of empirical curve fitting to different experimental data. It can be shown that the specific discharge ratios are equal if the exponent parameters of these options are equal. Comparison of specific discharge estimates from these two methods for a few tested headwater basins shows a linear relationship indicating a good agreement between the exponent parameter values.

RESULTS AND DISCUSSION

The tests described here have been performed on watersheds within the Arkansas–Red River basin in Oklahoma. The main reason is that this region has the longest archive of 4-km NEXRAD Stage III precipitation grids (NWS archive) derived from radar-gauge analyses (Seo & Breidenbach, 2002), and these rainfall estimates have been evaluated more thoroughly than those produced in other parts of the country. In addition, these basins have few complications such as snow, reservoirs, and complex topography such as mountainous areas.

A priori SAC-SMA parameter grids at the 4 km scale were generated over the entire region. These grids display a large variability of basin properties that affect runoff generation processes over this region. Initial routing parameter grids were derived using measured discharge data from only a few stations on the main stem of the Arkansas River. A parameter estimation procedure to derive values for upstream grid cells is applied sequentially from downstream to upstream gauges independently. At each estimation loop, grid cell parameters above a selected gauge, estimated previously from the next downstream gauge, are overwritten by new estimates from the selected upstream gauge. For analysis, these grids were then refined over a few selected headwater basins.

Two critical questions to consider in evaluating the potential benefits of distributed modelling are: (a) whether a distributed model can produce simulations that are comparable to or better than simulations from existing lumped models, and (b) whether it is possible to define a distributed parameter model calibration strategy that is robust across spatial scales. We consider the ability to produce simulations that are comparable to lumped results a positive result because there are other potential benefits from running a distributed model: the ability to simulate flows at small, ungauged outlets within a watershed, and the ability to incorporate future sources of spatial data are good examples.

Continuous discharge simulations at an hourly time step were generated for the 8-year period. HL-RMS results are evaluated against both hourly observed hydrographs and simulated hydrographs from a lumped version of the SAC-SMA model paired with a unit hydrograph routing procedure. First, "reference" results were generated using *a priori* estimates of water balance and routing parameters for both HL-RMS and a lumped application of SAC-SMA. Second, SAC-SMA parameters were calibrated in a lumped mode for all parent basins. *A priori* SAC-SMA parameter grids of HL-RMS were then adjusted as discussed above, using lumped calibration results for defined watersheds. Major adjustments to gridded parameter values were

Table 1 Accuracy statistics for selected headwater basins in the Arkansas–Red basins.

Watershed	Area (km^2)	HL-RMS simulations $*_{vol.}$	$*_{peak.}$	$*_{RMS}$	R	Lumped simulations $*_{vol.}$	$*_{peak.}$	$*_{RMS}$	R
Calibrated parameters:									
Blue	1232.8	25.0	***25.0***	***139.0***	***0.87***	***23.0***	35.0	141.0	0.86
Eldon	795.1	***16.4***	***25.7***	131.0	0.90	18.5	26.0	***114.5***	***0.92***
Tahlequah	2483.8	***11.3***	***20.5***	70.8	0.92	12.6	25.8	***64.9***	***0.94***
Watts	1644.6	***11.9***	***26.4***	86.5	0.92	12.9	30.2	***82.8***	***0.93***
Savoy*	432.8	***19.9***	52.2	***185.6***	***0.86***	20.9	***52.0***	196.7	0.85
Kansas*	284.9	23.8	***53.0***	***161.7***	***0.81***	***23.7***	55.8	189.6	0.73
Uncalibrated parameters:									
Blue	1232.8	38.0	***40.3***	187.0	0.81	***31.0***	42.8	***163.0***	***0.83***
Eldon	795.1	***27.4***	***45.3***	***160.0***	***0.84***	30.2	53.4	181.9	0.79
Tahlequah	2483.8	***13.4***	***19.2***	***84.5***	***0.90***	23.7	25.6	97.2	0.86
Watts	1644.6	***13.8***	***26.0***	***105.8***	***0.89***	23.1	30.5	109.9	0.87
Savoy	432.8	***22.4***	***49.8***	***194.2***	***0.85***	29.1	54.5	228.6	0.79
Kansas	284.9	***24.2***	***52.2***	***185.1***	***0.75***	26.9	57.1	221.1	0.62

Notes:

(a) Watersheds marked by * are nested basins (Savoy and Kansas are sub-basins of Tahlequah, Illinois River). These sub-basins actually were not calibrated; instead parent basin calibrated parameters were used in simulations.

(b) Values in bold italic mean that this statistic is better for distributed (lumped) simulations compared to lumped (distributed).

achieved by direct scaling from lumped calibrated parameters. Further minor adjustments of a few upper zone parameter scale factors were made using a manual trial-and-error process similar to an approach suggested by Vieux & Moreda (2002).

Table 1 presents statistical analyses of these simulations. Shown are two statistics for selected flood events (20–25 flood events for each basin were included in the analyses): percent of mean absolute error of flood runoff, $*_{vol}$, and flood peak, $*_{peak}$. Also presented are two overall statistics describing the entire simulation period: percent of root mean square error, $*_{RMS}$, and a correlation coefficient of hourly discharges, R. A few observations can be made about the results in Table 1.

While distributed and lumped simulations can both produce reasonable flood event simulations for calibrated watersheds, distributed model results are slightly better (compare columns 2–3 *vs* 6–7; note that better statistics are in bold italic). The distributed model benefit can also be seen in Fig. 1, which is a plot of flood peak errors for all selected floods from distributed and lumped runs over three basins. However, only the Blue River simulations show improvements over lumped simulations for the overall statistics of continuous runs (compare columns 4–5 *vs* 8–9). Other watersheds yield overall results comparable to lumped simulations with a slight decrease in accuracy.

It is important to note that simulations for two nested basins, Savoy and Kansas (not calibrated directly, but parameters from the calibration of the parent basin are applied instead), outperformed lumped simulations even more consistently than we would expect from a distributed model.

Consistent improvements from the uncalibrated parameter version of HL-RMS over uncalibrated lumped simulations are achieved for most watersheds (compare

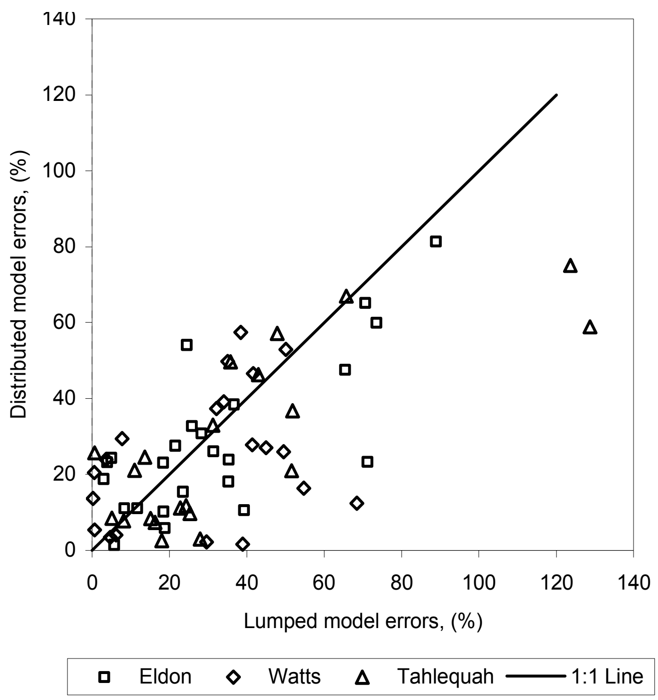

Fig. 1 Absolute flood peak errors (%) from distributed and lumped runs over three parent basins.

columns 2–5 *vs* 6–9 of the Uncalibrated parameters section of Table 1). It means that spatial patterns of *a priori* parameters and rainfall interact reasonably well, and as a result, the parameter scaling procedure used in this analysis may be a reasonable component in a multi-step calibration.

Tests suggest that HL-RMS provides a flexible framework for rainfall–runoff analysis and practical applications of distributed models, and that it is computationally feasible to run the system over large regions. The approach adds more practicality to the process of model parameterization, and facilitates an easier transition from current lumped model-based operational systems to more powerful distributed systems. Schemes developed to estimate distributed routing parameters based on local grid cell and basin outlet integrated properties produce reasonable results for the range of spatial scales without any calibration. Progress has been made in quantitatively estimating spatially-variable rainfall–runoff parameters by combining soil properties and lumped calibration results, but a robust method to calibrate the distributed parameter model remains to be developed.

REFERENCES

Beven, K. (1995) Linking parameters across scales: sub-grid parameterizations and scale dependent hydrological models. In: *Scale Issues in Hydrological Modelling*, (ed. by J. D. Kalma & M. Sivapalan), 263–282. Wiley, New York.

Beven, K. & Feyan, J. (eds) *The Future of Distributed Hydrological Modelling* (2002) Special Issue of *Hydrol. Processes* **16**(2) 169–574.

Burnash, R. J. C. (1995) The NWS river forecast system—catchment modelling. In: *Computer Models of Watershed Hydrology*, (ed. by V. P. Singh), 311–366. Water Resources Publications, Littleton, Colorado, USA.

Koren, V. I., Finnerty, B. D., Schaake, J. C., Smith, M. B., Seo, D. J. & Duan, Q. Y. (1999) Scale dependencies of hydrologic models to spatial variability of precipitation. *J. Hydrol.* **217**, 285–302.

Koren, V. I., Smith, M., Wang, D., Zhang, Z. (2000) Use of soil property data in the derivation of conceptual rainfall–runoff model parameters, 15th *Conference on Hydrology*, AMS, 103–106. Long Beach, California, USA

Koren, V., Smith & M., Duan, Q. (2002) Use of *a priori* parameter estimates in the derivation of spatially consistent parameter sets of rainfall–runoff models. In: *Calibration of Watershed Models, Water Science and Application 6* (ed. by Q. Duan, S. Sorooshian, H. Gupta, A. Rosseau & R. Turcotte), 239–254. Am. Geophys. U., Washington DC, USA.

Loague, K. (1990) R-5 Revisited 2. Reevaluation of a quasi-physically-based rainfall–runoff model with supplemental information. *Water Resour. Res.* **26**(5) 973–987.

Robinson, J. S. & Sivapalan M. (1995) Catchment-scale runoff generation model by aggregation and similarity analysis. In: *Scale Issues in Hydrological Modelling*, (ed. by J. D. Kalma & M. Sivapalan), 311–330. Wiley, New York, USA.

Seo, D.-J. & Breidenbach, J. P. (2002) Real-time correction of spatially nonuniform bias in radar rainfall data using rain gauge measurements. *J. Hydromet.* **1**, 222–240.

Smith, A. A. (1980) A generalized approach to kinematic flood routing. *J. Hydrol.* **45**, 71–89.

Vieux, B. E. & Moreda, F. G. (2002) Ordered physics-based parameter adjustment of a distributed model. In: *Calibration of Watershed Models, Water Science and Application 6* (ed. by Q. Duan, S. Sorooshian, H. Gupta, A. Rosseau & R. Turcotte), 267–281. Am. Geophys. U., Washington DC, USA.